

Plaquette 2 - Nuage de points et point moyen

Statistiques à deux variables — Terminale STMG

Corrections détaillées

Boîte à outils

Le nuage de points. On représente les couples $(x_i; y_i)$ dans un repère : x est la variable **explicative**, y la variable **expliquée**.

Le point moyen.

$$G(\bar{x}; \bar{y}) \quad \bar{x} = \frac{x_1 + \dots + x_n}{n}$$

La droite d'ajustement par les moindres carrés passe **toujours** par G .

Juger à l'œil.

Allure du nuage	Ajustement affine
points alignés ou presque	pertinent
forme en courbe	non pertinent, il faut changer de variable
nuage sans forme	non pertinent

Le coefficient de corrélation r .

Valeur de r	Lecture
proche de 1	forte liaison croissante
proche de -1	forte liaison décroissante
proche de 0	pas de liaison affine

En pratique, on considère l'ajustement acceptable quand $|r| > 0,95$.

⚠ Une forte corrélation ne prouve **aucune** causalité : elle décrit une tendance conjointe.

Le point moyen comme repère. Il est le « centre de gravité » du nuage : si le nuage est symétrique autour de lui, l'ajustement affine est souvent bon.

Correction de l'exercice 1 - Calculer un point moyen

— **Somme des x_i** : $1 + 2 + 3 + 4 + 5$ ✓

— **Total** : 15 ✓

— **Effectif** : 5 ✓

— **Moyenne** : $\frac{15}{5}$ ✓

— **Valeur** :

$$\bar{x} = 3$$

— **Somme des y_i , par paquets** : $42 + 48 = 90$ ✓

— **Puis** : $90 + 57 = 147$ ✓

— **Puis** : $147 + 63 + 70$ ✓

— **Total** : 280 ✓

— **Moyenne** : $\frac{280}{5}$ ✓

— **Valeur :**

$$\bar{y} = 56$$

— **Point moyen :**

$$G(3; 56)$$

— **Contrôle de vraisemblance :** 3 est au milieu de 1 et 5 ✓

— **Et :** 56 est entre 42 et 70 ✓

— **Remarque :** G n'est pas un point du nuage ici ✓

— **Vérification :** aucun couple ne vaut (3; 56) ✓

— **Le couple le plus proche :** (3; 57) ✓

— **Propriété à retenir :** la droite d'ajustement passe par G ✓

Réponse. $G(3; 56)$.

Correction de l'exercice 2 - Un nuage à décrire

— **Écarts en x :** 1 à chaque fois ✓

— **Écart de y , de 42 à 48 :** 6 ✓

— **De 48 à 57 :** 9 ✓

— **De 57 à 63 :** 6 ✓

— **De 63 à 70 :** 7 ✓

— **Ces écarts sont proches** ✓

— **Moyenne des écarts :** $\frac{28}{4}$ ✓

— **Valeur :** 7 ✓

— **Conclusion :**

ajustement affine pertinent

— **Justification :** les points sont presque alignés ✓

— **Pente approximative :** 7 ✓

— **Contrôle avec le point moyen :** la droite passe par $G(3; 56)$ ✓

— **Équation approchée :** $y = 7(x - 3) + 56$ ✓

— **Développement :** $y = 7x + 35$ ✓

— **Contrôle en $x = 1$:** 42 ✓

— **Contrôle en $x = 5$:** 70 ✓

— **Excellente concordance aux extrémités** ✓

— **Contrôle en $x = 3$:** 56, contre 57 observé ✓

— **Écart :** 1 ✓

— **Appréciation :** très faible ✓

Réponse. Oui : les écarts successifs (6, 9, 6, 7) sont proches, les points sont presque alignés.

Correction de l'exercice 3 - Deux nuages à comparer

a) **Critère usuel :** $|r| > 0,95$ ✓

— **Nuage A :** $|0,97| = 0,97$ ✓

— **Supérieur au seuil** ✓

— **Ajustement pertinent** ✓

— **Nuage B :** $|0,42| = 0,42$ ✓

- **Bien inférieur** ✓
- **Ajustement non pertinent** ✓
- **Nuage C** : $|-0,99| = 0,99$ ✓
- **Ajustement pertinent** ✓
- b) **Nuage A** : r positif ✓
- **Liaison** :
croissante
- **Lecture** : quand x augmente, y augmente ✓
- **Nuage C** : r négatif ✓
- **Liaison** :
décroissante
- **Lecture** : quand x augmente, y diminue ✓
- **Nuage B** : aucune liaison affine nette ✓
- **Ce que cela ne signifie pas** : qu'il n'y a aucun lien ✓
- **Cas possible** : une liaison en forme de courbe ✓
- **Exemple** : une parabole a un r proche de 0 ✓
- **Que faire alors** : tracer le nuage ✓
- **Et envisager un changement de variable** ✓
- **Rappel** : r ne mesure que la liaison **affine** ✓

Réponse. a) A et C · b) A croissante, C décroissante, B sans liaison affine.

Correction de l'exercice 4 - Point moyen d'une série chronologique

- **Valeurs de x_i** : 0 ; 1 ; 2 ; 3 ; 4 ; 5 ✓
- **Intérêt du changement** : des nombres plus petits ✓
- **Somme des x_i** : $0 + 1 + 2 + 3 + 4 + 5$ ✓
- **Total** : 15 ✓
- **Effectif** : 6 ✓
- **Moyenne** : $\frac{15}{6}$ ✓
- **Valeur** :
 $\bar{x} = 2,5$
- **Somme des y_i** : $45 + 52 = 97$ ✓
- **Puis** : $97 + 56 = 153$ ✓
- **Puis** : $153 + 63 = 216$ ✓
- **Puis** : $216 + 71 + 78$ ✓
- **Total** : 365 ✓
- **Moyenne** : $\frac{365}{6}$ ✓
- **Valeur** :
 $\bar{y} \approx 60,83$
- **Point moyen** : $G(2,5 ; 60,83)$ ✓
- **Année correspondante** : 2020,5 ✓
- **Interprétation** : le milieu de la période ✓
- **Contrôle** : 60,83 est entre 56 et 63 ✓

Réponse. $G(2,5 ; \approx 60,83)$, soit le milieu de la période.

Correction de l'exercice 5 - Lire une pente

a) **Nature du coefficient** : le coefficient directeur ✓

— **Interprétation** :

y augmente de 7,1 quand x augmente de 1

— **Précision** : en moyenne ✓

— **Signe** : positif ✓

— **Lecture** : la tendance est croissante ✓

b) **Nature** : l'ordonnée à l'origine ✓

— **C'est la valeur prédite pour $x = 0$** ✓

— **Valeur** :

$34,7$

— **Prudence** : $x = 0$ peut être hors du nuage ✓

— **Ici** : les x vont de 1 à 5 ✓

— **Il s'agit donc d'une extrapolation** ✓

c) **Substitution** : $7,1 \times 6 + 34,7$ ✓

— **Produit** : $42,6$ ✓

— **Somme** : $42,6 + 34,7$ ✓

— **Valeur** :

$y = 77,3$

— **Nature de la prévision** : une extrapolation ✓

— **Raison** : 6 dépasse le plus grand x observé ✓

— **Appréciation** : proche du nuage, donc raisonnable ✓

— **Prévision pour $x = 20$** : $176,7$ ✓

— **Appréciation** : très risquée ✓

— **Contrôle avec le point moyen** : $7,1 \times 3 + 34,7$ ✓

— **Valeur** : 56 ✓

— **Concordance avec $G(3; 56)$** ✓

Réponse. a) $+7,1$ par unité de x · b) la valeur prédite en $x = 0$ · c) $77,3$.

Correction de l'exercice 6 - Une liaison décroissante

a) **Somme des x_i** : $10 + 20 + 30 + 40 + 50$ ✓

— **Total** : 150 ✓

— **Moyenne** :

$\bar{x} = 30$

— **Somme des y_i** : $180 + 165 = 345$ ✓

— **Puis** : $345 + 152 = 497$ ✓

— **Puis** : $497 + 138 + 125$ ✓

— **Total** : 760 ✓

— **Moyenne** : $\frac{760}{5}$ ✓

— **Valeur** :

$\bar{y} = 152$

- **Point moyen** : $G(30; 152)$ ✓
- **Remarque** : ici G est un point du nuage ✓
- b) **Variation totale de y** : $125 - 180$ ✓
- **Valeur** : -55 ✓
- **Variation totale de x** : 40 ✓
- **Pente approximative** : $\frac{-55}{40}$ ✓
- **Valeur** : $-1,375$ ✓
- **Signe** : négatif ✓
- **Interprétation** : quand x augmente de 1, y baisse de $\approx 1,37$ ✓
- **Écarts successifs de y** : -15 ; -13 ; -14 ; -13 ✓
- **Ils sont très réguliers** ✓
- **Conclusion** : l'ajustement affine convient bien ✓
- **Coefficient de corrélation attendu** : proche de -1 ✓
- **Valeur réelle** : $\approx -0,9997$ ✓

Réponse. a) $G(30; 152)$ · b) pente $\approx -1,37$: liaison décroissante très régulière.

Correction de l'exercice 7 - Reconnaître une forme

- a) **Écarts successifs** : 10 ; 10 ; 10 ✓
- **Ils sont constants** ✓
- **Conclusion** :
parfaitement affine
- **Équation exacte** : $y = 10x$ ✓
- **Corrélation** : $r = 1$ ✓
- b) **Écarts successifs** : 2 ; 4 ; 8 ✓
- **Ils doublent** ✓
- **Ce n'est pas affine** ✓
- **Quotients successifs** : 2 ; 2 ; 2 ✓
- **Ils sont constants** ✓
- **Nature** :
modèle exponentiel
- **Que faire** : changer de variable ✓
- **Piste** : travailler sur $\log y$ ou ajuster $y = 2^x$ ✓
- c) **Écarts successifs** : 3 ; -1 ; 3 ✓
- **Ils ne sont pas constants** ✓
- **Ils changent même de signe** ✓
- **Tendance globale** : croissante ✓
- **Mais irrégulière** ✓
- **Conclusion** : ajustement affine **médiocre** ✓
- **Corrélation estimée** : autour de $0,8$ ✓
- **Décision** : chercher plus de données ✓
- **Ou accepter une prévision peu précise** ✓
- **Réflexe général** : regarder les **écarts** pour l'ajustement ✓
- **Et les quotients pour l'exponentiel** ✓

Réponse. a) oui, parfaitement · b) non, c'est exponentiel · c) médiocre, la série est irrégulière.

Correction de l'exercice 8 - Corrélation et causalité

a) **Valeur de r** : 0,96 ✓

— **Elle est élevée** ✓

— **Ce que cela prouve** : une liaison **statistique** ✓

— **Ce que cela ne prouve pas** :

aucune causalité

— **Réponse** : non ✓

b) **Variable cachée** : la chaleur ✓

— **Effet sur les glaces** : elles se vendent plus ✓

— **Effet sur la baignade** : il y a plus de baigneurs ✓

— **Conséquence** : plus de noyades ✓

— **Nom du phénomène** : une corrélation **fallacieuse** ✓

— **Vocabulaire** : la chaleur est une variable **confondante** ✓

— **Test de bon sens** : interdire les glaces réduirait-il les noyades ? ✓

— **Réponse** : non ✓

— **Preuve que la causalité est fausse** ✓

— **Autre exemple classique** : la taille et le vocabulaire chez l'enfant ✓

— **Variable cachée** : l'âge ✓

— **Ce que permet la corrélation** : **prévoir** ✓

— **Ce qu'elle ne permet pas** : **agir** ✓

— **Pour établir une causalité** : il faut une expérience contrôlée ✓

— **Règle à retenir** : corrélation $\neq$ causalité ✓

Réponse. Non : la chaleur explique les deux phénomènes. Corrélation n'est pas causalité.

Correction de l'exercice 9 - Prévoir avec le point moyen

a) **Propriété** : la droite passe par G ✓

— **Formule** : $y = a(x - \bar{x}) + \bar{y}$ ✓

— **Substitution** : $y = 2,5(x - 8) + 45$ ✓

— **Développement** : $2,5x - 20 + 45$ ✓

— **Équation** :

$$y = 2,5x + 25$$

— **Contrôle en $x = 8$** : $20 + 25 = 45$ ✓

— **Concordance avec G** ✓

b) **Substitution** : $2,5 \times 12 + 25$ ✓

— **Produit** : 30 ✓

— **Somme** : 55 ✓

— **Prévision** :

$$y = 55$$

— **Écart avec $\bar{y}$** : 10 ✓

- **Contrôle** : 4 unités de x à 2,5 ✓
- **Produit** : 10 ✓
- **Concordance** ✓
- **Prévision pour** $x = 0$: 25 ✓
- **Interprétation** : la valeur de base ✓
- **Prévision pour** $x = 20$: 75 ✓
- **Appréciation** : extrapolation lointaine ✓
- **Méthode rapide sans équation** : partir de G et compter les pas ✓
- **De 8 à 12** : 4 pas ✓
- **Gain** : $4 \times 2,5 = 10$ ✓
- **Prévision** : $45 + 10 = 55$ ✓

Réponse. a) $y = 2,5x + 25$ · b) $y = 55$.

Correction de l'exercice 10 - Interpolation ou extrapolation ?

a) **Position** : entre 2 et 12 ✓

— **Nature** :

interpolation

- **Fiabilité** : bonne ✓
- **Justification** : la tendance est observée sur cette plage ✓

b) **Position** : juste au-delà de 12 ✓

— **Nature** :

extrapolation

- **Fiabilité** : acceptable ✓
- **Justification** : on s'écarte peu du domaine observé ✓
- **Réserve** : la tendance peut déjà s'infléchir ✓

c) **Position** : très loin au-delà ✓

- **Nature** : extrapolation lointaine ✓
- **Fiabilité** :

faible

- **Justification** : rien ne garantit la tendance sur 28 unités de plus ✓
- **Exemples de ruptures possibles** : saturation du marché ✓
- **Ou** : changement de contexte économique ✓
- **Règle pratique** : n'extrapoler que de quelques pas ✓
- **Ordre de grandeur admis** : un tiers de l'amplitude observée ✓
- **Ici** : l'amplitude vaut 10 ✓
- **Extrapolation raisonnable** : jusqu'à $x \approx 15$ ✓
- **Conseil de rédaction** : toujours signaler l'extrapolation ✓
- **Et donner une réserve explicite** ✓

Réponse. a) interpolation, fiable · b) extrapolation proche, acceptable · c) extrapolation lointaine, peu fiable.

Correction de l'exercice 11 - Repérer une erreur

- **Erreur** : il confond le point moyen avec la valeur centrale ✓

— **Définition correcte** : la **moyenne** de chaque série ✓

— **Moyenne des x** : $\frac{15}{5}$ ✓

— **Valeur** : 3 ✓

— **Cette coordonnée est juste, par chance** ✓

— **Raison** : les x sont régulièrement espacés ✓

— **Moyenne des y** : $\frac{280}{5}$ ✓

— **Valeur** :

56

— **Sa réponse** : 57 ✓

— **Écart** : 1 ✓

— **Point moyen correct** :

$G(3; 56)$

— **Cas où l'erreur serait grave** : des x irréguliers ✓

— **Exemple** : $x = 1 ; 2 ; 3 ; 4 ; 20$ ✓

— **Valeur centrale** : 3 ✓

— **Moyenne** : 6 ✓

— **Écart** : 3 ✓

— **Autre cas grave** : une valeur aberrante en y ✓

— **Règle** : le point moyen se **calcule**, il ne se lit pas ✓

— **Vérification systématique** : il doit être dans le nuage ✓

— **Et la droite d'ajustement doit y passer** ✓

Réponse. $G(3; 56)$: le point moyen est la moyenne des deux séries, pas la valeur centrale.

Correction de l'exercice 12 - Deux séries chronologiques

a) **Série A, écarts** : 3 à chaque fois ✓

— **Nature** :

arithmétique

— **Raison** : les différences sont constantes ✓

— **Série B, écarts** : 6 ; 9 ; 13,5 ; 20,25 ✓

— **Ils ne sont pas constants** ✓

— **Quotients** : $\frac{18}{12} = 1,5$ ✓

— **Puis** : $\frac{27}{18} = 1,5$ ✓

— **Puis** : $\frac{40,5}{27} = 1,5$ ✓

— **Nature** :

géométrique de raison 1,5

b) **Série A** : les points sont exactement alignés ✓

— **Ajustement affine** : parfait ✓

— **Corrélation** : $r = 1$ ✓

— **Série B** : le nuage est **incurvé** ✓

— **Ajustement affine** : mauvais ✓

- **Solution** : changer de variable ✓
- **Piste** : poser $z = \log y$ ✓
- **Effet** : la série z devient arithmétique ✓
- **Équation de la droite pour A** : $y = 3x + 9$ avec $x = 1$ à 5 ✓
- **Contrôle en** $x = 1$: 12 ✓
- **Contrôle en** $x = 5$: 24 ✓
- **Prévision pour** $x = 8$ avec A : 33 ✓

Réponse. a) A arithmétique, B géométrique de raison 1,5 · b) seule A se prête à un ajustement affine.

Correction de l'exercice 13 - Une valeur aberrante

a) **Moyenne des** x : 3 ✓

- **Somme des** y : $20 + 23 = 43$ ✓
- **Puis** : $43 + 26 = 69$ ✓
- **Puis** : $69 + 29 + 80$ ✓
- **Total** : 178 ✓
- **Moyenne** : $\frac{178}{5}$ ✓
- **Valeur** :

$$\bar{y} = 35,6$$

— **Point moyen** : $G(3; 35,6)$ ✓

b) **Écarts sur les quatre premiers points** : 3 à chaque fois ✓

- **Dernier écart** : $80 - 29$ ✓
- **Valeur** : 51 ✓
- **Il est dix-sept fois plus grand** ✓
- **Diagnostic** : une valeur **aberrante** ✓
- **Effet sur** $\bar{y}$: elle la tire fortement vers le haut ✓
- **Sans cette valeur** : $\bar{y} = 24,5$ ✓
- **Écart** : 11,1 ✓
- **Position de G** : au-dessus de la plupart des points ✓
- **Conséquence sur l'ajustement** : la droite est faussée ✓
- **Ce qu'il faut faire** : vérifier la donnée ✓
- **Si c'est une erreur de saisie** : la corriger ✓
- **Si elle est réelle** : la signaler et l'expliquer ✓
- **Ne jamais la supprimer en silence** ✓

Réponse. a) $G(3; 35,6)$ · b) la valeur 80 est aberrante et fausse l'ajustement.

Correction de l'exercice 14 - Choisir la variable explicative

a) **Variable qu'on contrôle****** : la surface ✓

- **Variable qu'on cherche à prévoir****** : le chiffre d'affaires ✓
- **Règle** : l'explicative en abscisse ✓
- **Conclusion** :

$$x = \text{la surface}$$

- **Et** : y le chiffre d'affaires ✓
- **Raison pratique** : on décide d'une surface, pas d'un CA ✓

b) **Unité de x** : le mètre carré ✓

— **Unité de y** : l'euro ✓

— **Unité de la pente** : l'euro par mètre carré ✓

— **Interprétation** :

le CA gagné par m^2 supplémentaire

— **Nom courant en gestion** : le rendement au m^2 ✓

— **Exemple, pente** = 2400 : chaque m^2 rapporte 2400 € ✓

— **Interprétation de l'ordonnée à l'origine** : le CA d'un magasin de 0 m^2 ✓

— **Sens concret** : aucun ✓

— **Ce qu'elle représente vraiment** : un simple ajustement ✓

— **Si l'on inversait les axes** : la pente serait en m^2 par euro ✓

— **Ce ne serait pas la même droite** ✓

— **Raison** : les moindres carrés minimisent les écarts en y ✓

— **Conséquence** : l'ordre des variables compte ✓

Réponse. a) la surface · b) le chiffre d'affaires gagné par m^2 supplémentaire.

Correction de l'exercice 15 – Estimer une pente à la main

— **Première méthode, variation de y** : $336 - 250$ ✓

— **Valeur** : 86 ✓

— **Variation de x** : 5 ✓

— **Pente** : $\frac{86}{5}$ ✓

— **Valeur** :

17,2

— **Seconde méthode, écarts successifs** : 18 ; 17 ; 14 ; 19 ; 18 ✓

— **Somme** : 86 ✓

— **Nombre d'écarts** : 5 ✓

— **Moyenne** : 17,2 ✓

— **Concordance** : les deux méthodes coïncident ✓

— **Explication** : la somme des écarts est la variation totale ✓

— **Valeur exacte par les moindres carrés** : $\approx 16,97$ ✓

— **Écart avec notre estimation** : $\approx 0,23$ ✓

— **Appréciation** : très proche ✓

— **Pourquoi un léger écart** : les moindres carrés pondèrent tous les points ✓

— **Point moyen** : $\bar{x} = 2,5$ ✓

— **Somme des y** : 1756 ✓

— **Moyenne** : $\approx 292,67$ ✓

— **Équation approchée** : $y = 17,2(x - 2,5) + 292,67$ ✓

— **Développement** : $y = 17,2x + 249,67$ ✓

— **Équation exacte** : $y \approx 16,97x + 250,24$ ✓

Réponse. Les deux méthodes donnent 17,2, contre $\approx 16,97$ par les moindres carrés.

Correction de l'exercice 16 – Synthèse : un nuage complet

a) **Somme des x** : 21 ✓

— **Effectif** : 6 ✓

— **Moyenne** :

$$\bar{x} = 3,5$$

— **Somme des y** : $12 + 15 = 27$ ✓

— **Puis** : $27 + 19 = 46$ ✓

— **Puis** : $46 + 22 = 68$ ✓

— **Puis** : $68 + 26 + 30$ ✓

— **Total** : 124 ✓

— **Moyenne** : $\frac{124}{6}$ ✓

— **Valeur** :

$$\bar{y} \approx 20,67$$

b) **Variation totale de y** : $30 - 12$ ✓

— **Valeur** : 18 ✓

— **Variation de x** : 5 ✓

— **Pente** : $\frac{18}{5}$ ✓

— **Valeur** :

$$3,6$$

— **Contrôle des écarts** : 3 ; 4 ; 3 ; 4 ; 4 ✓

— **Moyenne** : 3,6 ✓

— **Concordance** ✓

c) **Formule** : $y = 3,6(x - 3,5) + 20,67$ ✓

— **Développement** : $3,6x - 12,6 + 20,67$ ✓

— **Équation** :

$$y = 3,6x + 8,07$$

— **Prévision en x = 8** : $28,8 + 8,07$ ✓

— **Valeur** :

$$\approx 36,87$$

— **Nature** : une extrapolation ✓

— **Appréciation** : 8 est proche de 6, prévision raisonnable ✓

— **Corrélation de la série** : $\approx 0,9988$ ✓

— **Conclusion** : l'ajustement est excellent ✓

Réponse. a) $G(3,5; \approx 20,67)$ · b) 3,6 · c) $y = 3,6x + 8,07$ et $y(8) \approx 36,87$.